Proposed documentation may be developed by the ATSC, by member organizations of the JCIC, or by existing standards committees. The ATSC was established recognizing that the prompt, efficient and effective development of a coordinated set of national standards is essential to the future development of domestic television services.

On June 5, 1992 ATSC provided information to the Federal Communications Commission (FCC) outlining proposed industry actions to fully document the advanced television system standard. The FCC has recognized the importance of prompt disclosure of the system technical specifications to the mass production of advanced television system professional and consumer equipment in a timely fashion. The FCC has further noted its appreciation of the diligence with which the ATSC and the other groups participating in the standardization are pursuing these matters.[2]

Supporting this activity, the ATSC Executive Committee requested that the T3/S1 Specialist Group on Macro Systems Approach meet and suggest which portions of an advanced television system broadcasting standard might require action by the FCC and which portions should be voluntary.

Subsequently, T3/S1 held meetings and developed recommendations in two areas:

1. Principles upon which documentation of the advanced television system should be based; and
2. A list of characteristics of an advanced television system that should be documented.

The list tentatively identified the industry group(s) that would provide the documentation information and the document where the information would likely appear.

The recommendations developed by the T3/S1 Specialist Group were modified by T3 to accommodate information and knowledge about advanced television systems developed in the period since June 1992. Some of the modifications to the recommendations ensued from the formation of the Grand Alliance. The modified guidelines were approved at the March 31, 1994 meeting of the T3 Technology Group on Distribution and are described in Section 4.5.

4.2 Advisory Committee on Advanced Television Service (ACATS)

A "Petition for Notice of Inquiry" was filed with the FCC on February 21, 1987 by 58 broadcasting organizations and companies requesting that the Commission initiate a proceeding to explore the issues arising from the introduction of advanced television technologies and their possible impact on the television broadcasting service. At that time, it was generally believed that High Definition Television (HDTV) could not be broadcast using 6 MHz terrestrial broadcasting channels. The broadcasting organizations were concerned that the alternative media would be able to deliver HDTV to the viewing public placing terrestrial broadcasting at a severe disadvantage.

[2] FCC 92-438, MM Docket No. 87-268, "Memorandum Opinion and Order/Third Report and Order/Third Further Notice of Proposed Rule Making," Adopted: September 17, 1992, pp. 59-60.

The FCC agreed that this was a subject of utmost importance and initiated a proceeding (MM Docket No. 87-268) to consider the technical and public policy issues of advanced television systems. The Advisory Committee on Advanced Television Service was empaneled by the Federal Communications Commission in 1987 with Richard E. Wiley as chairman to develop information that would assist the FCC in establishing an advanced television standard for the United States. The objective given to the Advisory Committee in its Charter by the FCC was:

> "The Committee will advise the Federal Communications Commission on the facts and circumstances regarding advanced television systems for Commission consideration of technical and public policy issues. In the event that the Commission decides that adoption of some form of advanced broadcast television is in the public interest, the Committee would also recommend policies, standards and regulations that would facilitate the orderly and timely introduction of advanced television services in the United States."

The Advisory Committee established a series of subgroups to study the various issues concerning services, technical parameters, and testing mechanisms required to establish an Advanced television system standard. The Advisory Committee also established a system evaluation, test and analysis process that began with over twenty proposed systems, reducing them to four final systems for consideration.

4.3 Digital HDTV Grand Alliance (Grand Alliance)

On May 24, 1993 the three groups that had developed the four final digital systems agreed to produce a single, best-of-the best system to propose as the standard. The three groups (AT&T and Zenith Electronics Corporation; General Instrument Corporation and the Massachusetts Institute of Technology; and Philips Consumer Electronics, Thomson Consumer Electronics, and the David Sarnoff Research Center) have been working together as the "Digital HDTV Grand Alliance." The system described in this Standard is based on the Digital HDTV Grand Alliance proposal to the Advisory Committee.

4.4 Organization for documenting the Digital Television Standard

The ATSC Executive Committee assigned the work of documenting the advanced television system standards to T3 specialist groups dividing the work into five areas of interest: **Video** (including input signal format and source coding), **Audio** (including input signal format and source coding), **Transport** (including data multiplex and channel coding), **RF/Transmission,** (including the modulation subsystem) and **Receiver** characteristics. A steering committee consisting of the chairs of the five specialist groups, the chair and vice-chairs of T3, and liaison among the ATSC, the FCC, and ACATS was established to coordinate the development of the documents. The members of the steering committee and areas of interest were as follows:

Stanley Baron	T3 chair
Jules Cohen	T3 vice-chair
Brian James	T3 vice-chair

Larry Pearlstein	T3/S6 (Video systems characteristics), chair
Graham S. Stubbs	T3/S7 (Audio systems characteristics), chair
Bernard J. Lechner	T3/S8 (Service multiplex and transport systems characteristics), chair
Lynn D. Claudy	T3/S9 (RF/Transmission systems characteristics), chair
Werner F. Wedam	T3/S10 (Receiver characteristics), chair
Robert M. Rast	Grand Alliance facilitator
Robert Hopkins	ATSC
Robert M. Bromery	FCC Office of Engineering and Technology
Gordon Godfrey	FCC Mass Media Bureau
Paul E. Misener	ACATS

4.5 Principles for documenting the Digital Television Standard

T3 adopted the following principles for documenting the advanced television system standard:

1. The Grand Alliance was recognized as the principal supplier of information for documenting the advanced television system, supported by the ATSC and others. Other organizations seen as suppliers of information: EIA, FCC, IEEE, MPEG, NCTA, and SMPTE.
2. The Grand Alliance was encouraged to begin drafting the essential elements of system details as soon as possible to avoid delays in producing the advanced television system documentation.
3. FCC requirements for the advanced television system standard were to be obtained as soon as possible.
4. Complete functional system details (permitting those skilled in the art to construct a working system) were to be made publicly available.
5. Protection of any intellectual property made public must be by patent or copyright as appropriate.
6. The advanced television system documentation shall include the necessary system information such that audio and video encoders may be manufactured to deliver the system's full demonstrated performance quality.
7. The advanced television system documentation shall point to existing standards, recommended practices or guideline documents. These documents shall be referenced in one of two ways as deemed appropriate for the application. In the first instance, a specific revision shall be specified where review of changes to the referenced document is required before changes might be incorporated into the advanced television system document. The second instance references the document without specificity to revision and allows any changes to the referenced documents to be automatically incorporated.

8. System specifications shall explain how future, compatible improvements may be achieved.
9. As ongoing improvements take place in the advanced television system, manufacturers of encoders and decoders should coordinate their efforts to insure compatibility.
10. The advanced television system standard must support backward compatibility of future improvements with all generations of advanced television system receivers and inherently support production of low cost receivers (not withstanding that cost reduction through reduced performance quality may also be used to achieve inexpensive products).
11. The advanced television system standard should not foreclose flexibility in implementing advanced television system receivers at different price and performance levels.
12. The advanced television system standard should not foreclose flexibility in implementing program services or in data stream modification or insertion of data packets by down-stream (local) service providers.
13. The advanced television system documentation shall address interoperability with non-broadcast delivery systems including cable.
14. The advanced television system standard shall identify critical system parameters and shall provide information as to the range of acceptable values, the method of measurement, and the location in the system where measurement takes place.

5. SYSTEM OVERVIEW

5.1 Objectives

The Digital Television Standard describes a system designed to transmit high quality video and audio and ancillary data over a single 6 MHz channel. The system can deliver reliably about 19 Mbps of throughput in a 6 MHz terrestrial broadcasting channel and about 38 Mbps of throughput in a 6 MHz cable television channel. This means that encoding a video source whose resolution can be as high as five times that of conventional television (NTSC) resolution requires a bit rate reduction by a factor of 50 or higher. To achieve this bit rate reduction, the system is designed to be efficient in utilizing available channel capacity by exploiting complex video and audio compression technology.

The objective is to maximize the information passed through the data channel by minimizing the amount of data required to represent the video image sequence and its associated audio. The objective is to represent the video, audio, and data sources with as few bits as possible while preserving the level of quality required for the given application.

Although the RF/Transmission subsystems described in this Standard are designed specifically for terrestrial and cable applications, the objective is that the video, audio, and service multiplex/transport subsystems be useful in other applications.

5.2 System block diagram

A basic block diagram representation of the system is shown in Figure 5.1. This representation is based on one adopted by the International Telecommunication Union, Radiocommunication Sector (ITU-R), Task Group 11/3 (Digital Terrestrial Television Broadcasting). According to this model, the digital television system can be seen to consist of three subsystems.[3]

1. Source coding and compression,
2. Service multiplex and transport, and
3. RF/Transmission.

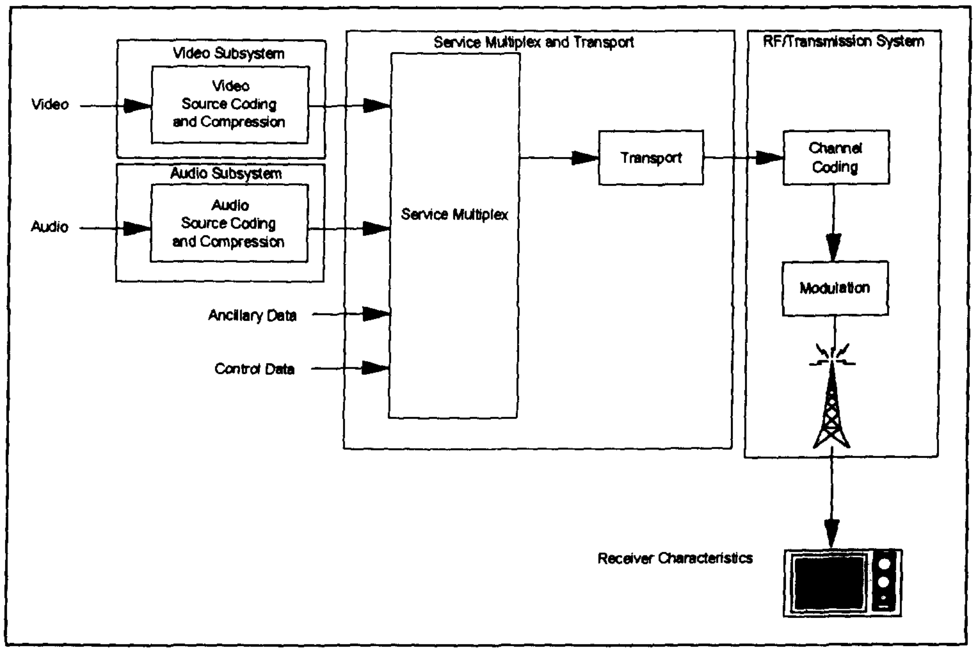

Figure 5.1. ITU-R digital terrestrial television broadcasting model.

"Source coding and compression" refers to the bit rate reduction methods, also known as data compression, appropriate for application to the video, audio, and ancillary digital data streams. The term "ancillary data" includes control data, conditional access control data, and data associated with the program audio and video services, such as closed captioning. "Ancillary data" can also refer to independent program services. The purpose of the coder is to minimize the number of bits needed to represent the audio and video information. The digital television system employs the MPEG-2 video stream syntax for the coding of video and the Digital Audio Compression (AC-3) Standard for the coding of audio.

"Service multiplex and transport" refers to the means of dividing the digital data stream into "packets" of information, the means of uniquely identifying each packet or packet type,

[3] ITU-R Document TG11/3-2, "Outline of Work for Task Group 11/3, Digital Terrestrial Television Broadcasting," June 30, 1992.

and the appropriate methods of multiplexing video data stream packets, audio data stream packets, and ancillary data stream packets into a single data stream. In developing the transport mechanism, interoperability among digital media, such as terrestrial broadcasting, cable distribution, satellite distribution, recording media, and computer interfaces, was a prime consideration. The digital television system employs the MPEG-2 transport stream syntax for the packetization and multiplexing of video, audio, and data signals for digital broadcasting systems.[4] The MPEG-2 transport stream syntax was developed for applications where channel bandwidth or recording media capacity is limited and the requirement for an efficient transport mechanism is paramount. It was designed also to facilitate interoperability with the ATM transport mechanism.

"RF/Transmission" refers to channel coding and modulation. The channel coder takes the data bit stream and adds additional information that can be used by the receiver to reconstruct the data from the received signal which, due to transmission impairments, may not accurately represent the transmitted signal. The modulation (or physical layer) uses the digital data stream information to modulate the transmitted signal. The modulation subsystem offers two modes: a terrestrial broadcast mode (8 VSB), and a high data rate mode (16 VSB).

Figure 5.2 illustrates a high level view of encoding equipment. This view is not intended to be complete, but is used to illustrate the relationship of various clock frequencies within the encoder. There are two domains within the encoder where a set of frequencies are related, the source coding domain and the channel coding domain.

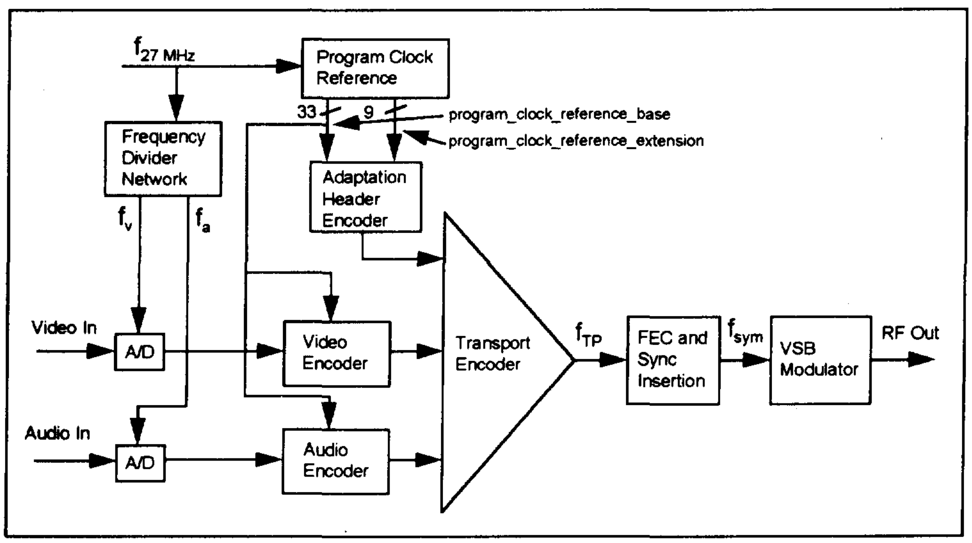

Figure 5.2. High level view of encoding equipment.

The source coding domain, represented schematically by the video, audio and transport encoders, uses a family of frequencies which are based on a 27 MHz clock (f_{27MHz}). This clock is used to generate a 42-bit sample of the frequency which is

[4] Chairman, ITU-R Task Group 11/3, "Report of the Second Meeting of ITU-R Task Group 11/3, Geneva, October 13-19, 1993," January 5, 1994, p. 40.

partitioned into two parts defined by the MPEG-2 specification. These are the 33-bit program_clock_reference_base and the 9-bit program_clock_reference_extension. The former is equivalent to a sample of a 90 kHz clock which is locked in frequency to the 27 MHz clock, and is used by the audio and video source encoders when encoding the presentation time stamp (PTS) and the decode time stamp (DTS). The audio and video sampling clocks, f_a and f_v respectively, must be frequency-locked to the 27 MHz clock. This can be expressed as the requirement that there exist two pairs of integers, (n_a, m_a) and (n_v, m_v), such that:

$$f_a = \left(\frac{n_a}{m_a}\right) x\ 27MHz$$

and

$$f_v = \left(\frac{n_v}{m_v}\right) x\ 27MHz$$

The channel coding domain is represented by the FEC/Sync Insertion subsystem and the VSB modulator. The relevant frequencies in this domain are the VSB symbol frequency (f_{sym})and the frequency of the transport stream (f_{TP}) which is the frequency of transmission of the encoded transport stream. These two frequencies must be locked, having the relation:

$$f_{TP} = 2\ x \left(\frac{188}{208}\right)\left(\frac{312}{313}\right) f_{sym}$$

The signals in the two domains are not required to be frequency-locked to each other, and in many implementations will operate asynchronously. In such systems, the frequency drift can necessitate the occasional insertion or deletion of a NULL packet from within the transport stream, thereby accommodating the frequency disparity.

The annexes that follow consider the characteristics of the subsystems necessary to accommodate the services envisioned.

ANNEX A

(Normative)

VIDEO SYSTEMS CHARACTERISTICS

1. SCOPE

This Annex describes the characteristics of the video subsystem of the Digital Television Standard. The input formats and bit stream characteristics are described in separate sections.

2. REFERENCES

2.1 Normative references

The following documents contain provisions which, through reference in this text, constitute provisions of this standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreement based on this standard are encouraged to investigate the possibility of applying the most recent editions of the documents listed below.

ISO/IEC IS 13818-1, International Standard (1994), *MPEG-2 Systems.*

ISO/IEC IS 13818-2, International Standard (1994), *MPEG-2 Video.*

2.2 Informative references

SMPTE 274M (1995), *Standard for television, 1920 x 1080 Scanning and Interface.*

SMPTE S17.392 (1995), *Proposed Standard for television, 1280 x 720 Scanning and Interface.*

ITU-R BT.601-4 (1994), *Encoding parameters of digital television for studios.*

3. COMPLIANCE NOTATION

As used in this document, *"shall"* or *"will"* denotes a mandatory provision of the standard. *"Should"* denotes a provision that is recommended but not mandatory. *"May"* denotes a feature whose presence does not preclude compliance, that may or may not be present at the option of the implementor.

4. POSSIBLE VIDEO INPUTS

While not required by this standard, there are certain television production standards, shown in Table 1, that define video formats that relate to compression formats specified by this standard.

Table 1 Standardized Video Input Formats

Video standard	Active lines	Active samples/ line
SMPTE 274M	1080	1920
SMPTE S17.392	720	1280
ITU-R BT.601-4	483	720

The compression formats may be derived from one or more appropriate video input formats. It may be anticipated that additional video production standards will be developed in the future that extend the number of possible input formats.

5. SOURCE CODING SPECIFICATION

The ATV video compression algorithm shall conform to the Main Profile syntax of ISO/IEC 13818-2. The allowable parameters shall be bounded by the upper limits specified for the Main Profile at High Level.[1] Additionally, ATV bit streams shall meet the constraints and specifications described in Sections 5.1 and 5.2.

5.1 Constraints with respect to ISO/IEC 13818-2 Main Profile

The following tables list the allowed values for each of the ISO/IEC 13818-2 syntactic elements which are restricted beyond the limits imposed by MP@HL.

In these tables conventional numbers denote decimal values, numbers preceded by **0x** are to be interpreted as hexadecimal values and numbers within single quotes (e.g., '10010100') are to be interpreted as a string of binary digits.

5.1.1 Sequence header constraints

Table 2 identifies parameters in the sequence header of a bit stream that shall be constrained by the video subsystem and lists the allowed values for each.

Table 2 Sequence Header Constraints

Sequence header syntactic element	Allowed value
horizontal_size_value	see Table 3
vertical_size_value	see Table 3
aspect_ratio_information	see Table 3
frame_rate_code	see Table 3
bit_rate_value (≤ 19.4 Mbps)	≤ 48500
bit_rate_value (≤ 38.8 Mbps)	≤ 97000
vbv_buffer_size_value	≤ 488

The allowable values for the field bit_rate_value are application dependent. In the primary application of terrestrial broadcast, this field shall correspond to a bit rate which is less than or equal to 19.4 Mbps. In the high data rate mode, the corresponding bit rate is less than or equal to 38.8 Mbps.

[1] See ISO/IEC 13818-2, Section 8 for more information regarding profiles and levels.

5.1.2 Compression format constraints

Table 3 lists the allowed compression formats.

Table 3 Compression Format Constraints

vertical_size_ value	horizontal_size_ value	aspect_ratio_ information	frame_rate_ code	progressive_ sequence
1080[2]	1920	1,3	1,2,4,5	1
			4,5	0
720	1280	1,3	1,2,4,5,7,8	1
480	704	2,3	1,2,4,5,7,8	1
			4,5	0
	640	1,2	1,2,4,5,7,8	1
			4,5	0

Legend for MPEG-2 coded values in Table 3
aspect_ratio_information 1 = square samples 2 = 4:3 display aspect ratio 3 = 16:9 display aspect ratio
frame_rate_code 1 = 23.976 Hz 2 = 24 Hz 4 = 29.97 Hz 5 = 30 Hz 7 = 59.94 Hz 8 = 60 Hz
progressive_sequence 0 = interlaced scan 1 = progressive scan

5.1.3 Sequence extension constraints

Table 4 identifies parameters in the sequence extension part of a bit stream that shall be constrained by the video subsystem and lists the allowed values for each. A sequence_extension structure is required to be present after every sequence_header structure.

Table 4 Sequence Extension Constraints

Sequence extension syntactic element	Allowed values
progressive_sequence	see Table 3
profile_and_level_indication	see Note
chroma_format	'01'
horizontal_size_extension	'00'
vertical_size_extension	'00'
bit_rate_extension	'0000 0000 0000'
vbv_buffer_size_extension	'0000 0000'
frame_rate_extension_n	'00'
frame_rate_extension_d	'0000 0'

Note: The profile_and_level_indication field shall indicate the lowest profile and level defined in ISO/IEC 13818-2, Section 8, that is consistent with the parameters of the video elementary stream.

[2] Note that 1088 lines are actually coded in order to satisfy the MPEG-2 requirement that the coded vertical size be a multiple of 16 (progressive scan) or 32 (interlaced scan).

5.1.4 Sequence display extension constraints

Table 5 identifies parameters in the sequence display extension part of a bit stream that shall be constrained by the video subsystem and lists the allowed values for each.

Table 5 Sequence Display Extension Constraints

Sequence display extension syntactic element	Allowed values
video_format	'000'

The preferred and default values for color_primaries, transfer_characteristics, and matrix_coefficients are defined to be SMPTE 274M[3] (value 0x01 in all three cases). While all values described by MPEG-2 are allowed in the transmitted bit stream, it is noted that SMPTE 170M values (0x06 in all three cases) will be the most likely alternate in common use.

5.1.5 Picture header constraints

In all cases other than when vbv_delay has the value 0xFFFF, the value of vbv_delay shall be constrained as follows:

vbv_delay $\leq$ 45000

5.2 Bit stream specifications beyond MPEG-2

This section covers the extension and user data part of the video syntax. These data are inserted at the sequence, GOP, and picture level. The syntax used for the insertion of closed captioning in picture user data is described.[4]

5.2.1 Picture extension and user data syntax

Table 6 describes the syntax used for picture extension and user data.

Table 6 Picture Extension and User Data Syntax

	No. of bits	Mnemonic
extension_and_user_data(2) {		
while ((nextbits( ) == extension_start_code) || (nextbits() == user_data_start_code)) {		
if (nextbits()== extension_start_code)		
extension_data(2)		
if (nextbits() == user_data_start_code)		
user_data(2)		
}		
}		

[3] At some point in the future, the color gamut may be extended by allowing negative values of RGB and defining the transfer characteristics for negative RGB values.

[4] In order to decode the user data, the decoder should properly recognize the 32-bit ATSC registration identifier at the PSI stream level (see ISO/IEC 13818-1).

5.2.2 Picture user data syntax

Table 7 describes the picture user data syntax.

Table 7 Picture User Data Syntax[5]

	No. of bits	Mnemonic
user_data( ) {		
user_data_start_code	32	bslbf
ATSC_identifier	32	bslbf
user_data_type_code	8	uimsbf
if (user_data_type_code == '0x03') {		
process_em_data_flag	1	bslbf
process_cc_data_flag	1	bslbf
additional_data_flag	1	bslbf
cc_count	5	uimsbf
em_data	8	bslbf
for (i=0 ; i < cc_count ; i++) {		
marker_bits	5	'1111 1'
cc_valid	1	bslbf
cc_type	2	bslbf
cc_data_1	8	bslbf
cc_data_2	8	bslbf
}		
marker_bits	8	'1111 1111'
if (additional_data_flag) {		
while(nextbits() != '0000 0000 0000 0000 0000 0001') {		
additional_user_data	8	
}		
}		
}		
next_start_code()		
}		

5.2.3 Picture user data semantics

user_data_start_code — This is set to 0x0000 01B2.

ATSC_identifier — This is a 32 bit code that indicates that the video user data conforms to this specification. The value ATSC_identifier shall be 0x4741 3934.

user_data_type_code — The 8-bit code is set to 0x03.

[5] Shaded cells in this table indicate syntactic and semantic additions to the ISO/IEC 13818-2 standard.

process_em_data_flag — This flag is set to indicate whether it is necessary to process the em_data. If it is set to 1, the em_data has to be parsed and its meaning has to be processed. When it is set to 0, the em_data can be discarded.

process_cc_data_flag — This flag is set to indicate whether it is necessary to process the cc_data. If it is set to 1, the cc_data has to be parsed and its meaning has to be processed. When it is set to 0, the cc_data can be discarded.

additional_data_flag — This flag is set to 1 to indicate the presence of additional user data.

cc_count — This 5-bit integer indicates the number of closed caption constructs following this field. It can have values 0 through 31. The value of cc_count shall be set according to the frame rate and coded picture structure (field or frame) such that a fixed bandwidth of 9600 bits per second is maintained for the closed caption payload data. Sixteen (16) bits of closed caption payload data are carried in each pair of the fields cc_data_1 and cc_data_2.

em_data — Eight bits for representing emergency message.[6]

cc_valid — This flag is set to '1' to indicate that the two closed caption data bytes that follow are valid. If set to '0' the two data bytes are invalid.

cc_type — Denotes the type of the two closed caption data bytes that follow.[7]

cc_data_1 — The first byte of a closed caption data pair.

cc_data_2 — The second byte of a closed caption data pair.

additional_user_data — Any further demand for picture user data could be met by defining this part of the bit stream.

[6] Syntax and semantics to be specified by EIA.

[7] EIA, *Recommended Practice for Advanced Television Closed Captioning*, draft, July 1, 1994.

ANNEX B

(Normative)

AUDIO SYSTEMS CHARACTERISTICS

1. SCOPE

This Annex describes the audio system characteristics and normative specifications of the Digital Television Standard.

2. NORMATIVE REFERENCES

The following documents contain provisions which in whole or part, through reference in this text, constitute provisions of this standard. At the time of publication, the editions indicated were valid. All standards are subject to revision and amendment, and parties to agreement based on this standard are encouraged to investigate the possibility of applying the most recent editions of the documents listed below.

ATSC Standard A/52 (1995), *Digital Audio Compression (AC-3).*

AES 3-1992 (ANSI S4.40-1992), *AES Recommended Practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data.*

ANSI S1.4-1983, *Specification for Sound Level Meters.*

IEC 651 (1979), *Sound Level Meters.*

IEC 804 (1985), Amendment 1 (1989) *Integrating/Averaging Sound Level Meters.*

3. COMPLIANCE NOTATION

As used in this document, *"shall"* or *"will"* denotes a mandatory provision of the standard. *"Should"* denotes a provision that is recommended but not mandatory. *"May"* denotes a feature whose presence does not preclude compliance, that may or may not be present at the option of the implementor.

4. SYSTEM OVERVIEW

As illustrated in Figure 1, the audio subsystem comprises the audio encoding/decoding function and resides between the audio inputs/outputs and the transport subsystem. The audio encoder(s) is (are) responsible for generating the audio elementary stream(s) which are encoded representations of the baseband audio input signals. At the receiver, the audio subsystem is responsible for decoding the audio elementary stream(s) back into baseband audio.

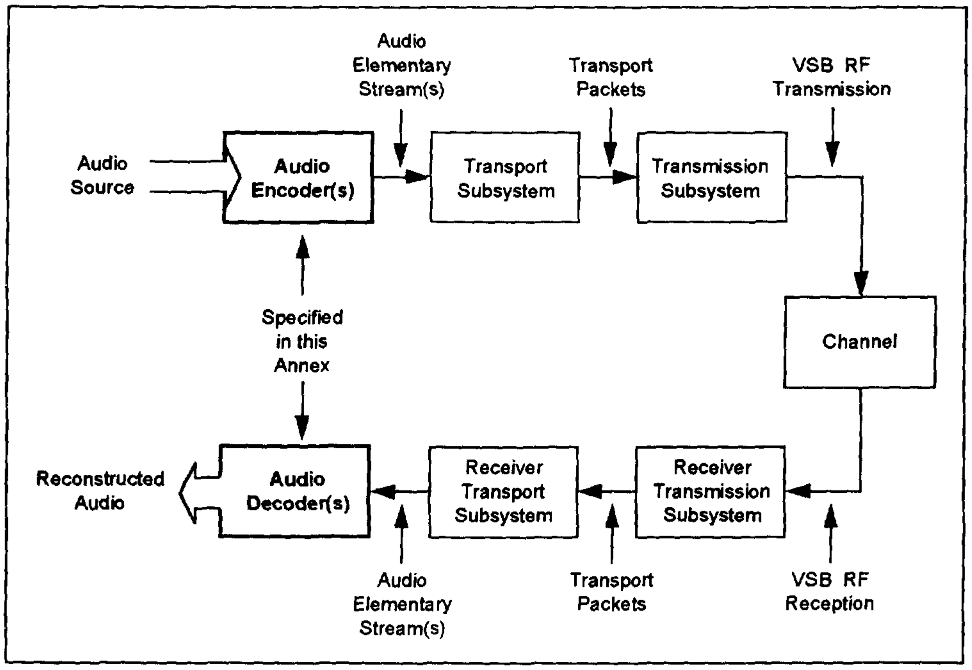

Figure 1. Audio subsystem in the digital television system.

5. SPECIFICATION

This Section forms the normative specification of the audio system. The audio compression system conforms with the Digital Audio Compression (AC-3) Standard, subject to the constraints outlined in this Section.

5.1 Constraints with respect to ATSC Standard A/52

The digital television audio coding system is based on the Digital Audio Compression (AC-3) Standard specified in the body of ATSC Doc. A/52 (the annexes are not included). Constraints on the system are shown in Table 1 which shows permitted values of certain syntactical elements. These constraints are described in Sections 5.2 - 5.4.

Table 1 Audio Constraints

AC-3 syntactical element	Comment	Allowed value
fscod	Indicates sampling rate	'00' (indicates 48 kHz)
frmsizecod	Main audio service or associated audio service containing all necessary program elements	≤ '011100' (indicates ≤ 384 kbps)
frmsizecod	Single channel associated service containing a single program element	≤ '010000' (indicates ≤ 128 kbps)
frmsizecod	Two channel dialogue associated service	≤ '010100' (indicates ≤ 192 kbps)
(frmsizecod)	Combined bit rate of a main and an associated service intended to be simultaneously decoded	(total ≤ 512 kbps)
acmod	Indicates number of channels	≥ '001'

5.2 Sampling frequency

The system conveys digital audio sampled at a frequency of 48 kHz, locked to the 27 MHz system clock. The 48 kHz audio sampling clock is defined as:

(1) 48 kHz audio sample rate = (2 ÷1125) × (27 MHz system clock)

If analog signal inputs are employed, the A/D converters should sample at 48 kHz. If digital inputs are employed, the input sampling rate shall be 48 kHz, or the audio encoder shall contain sampling rate converters which convert the sampling rate to 48 kHz.

5.3 Bit rate

A main audio service, or an associated audio service which is a complete service (containing all necessary program elements) shall be encoded at a bit rate less than or equal to 384 kbps. A single channel associated service containing a single program element shall be encoded at a bit rate less than or equal to 128 kbps. A two channel associated service containing only dialogue shall be encoded at a bit rate less than or equal to 192 kbps. The combined bit rate of a main service and an associated service which are intended to be decoded simultaneously shall be less than or equal to 512 kbps.

5.4 Audio coding modes

Audio services shall be encoded using any of the audio coding modes specified in A/52, with the exception of the 1+1 mode. The value of acmod in the AC-3 bit stream shall have a value in the range of 1-7, with the value 0 prohibited.

5.5 Dialogue level

The value of the dialnorm parameter in the AC-3 elementary bit stream shall indicate the level of average spoken dialogue within the encoded audio program. Dialogue level may be measured by means of an "A" weighted integrated measurement (LAeq). (Receivers use the value of dialnorm to adjust the reproduced audio level so as to normalize the dialogue level.)

5.6 Dynamic range compression

Each encoded audio block may contain a dynamic range control word (dynrng) which is used by decoders (by default) to alter the level of the reproduced audio. The control words allow the decoded signal level to be increased or decreased by up to 24 dB. In general, elementary streams may have dynamic range control words inserted or modified without affecting the encoded audio. When it is necessary to alter the dynamic range of audio programs which are broadcast, the dynamic range control word should be used.

6. MAIN AND ASSOCIATED SERVICES

6.1 Overview

An AC-3 elementary stream contains the encoded representation of a single audio service. Multiple audio services are provided by multiple elementary streams. Each elementary stream is conveyed by the transport multiplex with a unique PID. There are a number of audio service types which may (individually) be coded into each elementary stream. Each AC-3 elementary stream is tagged as to its service type using the bsmod bit field. There are two types of *main service* and six types of *associated service*. Each associated service may be tagged (in the AC-3 audio descriptor in the transport PSI data) as being associated with one or more main audio services. Each AC-3 elementary stream may also be tagged with a language code.

Associated services may contain complete program mixes, or may contain only a single program element. Associated services which are complete mixes may be decoded and used as is. They are identified by the full_svc bit in the AC-3 descriptor (see A/52, Annex A). Associated services which contain only a single program element are intended to be combined with the program elements from a main audio service.

This Section specifies the meaning and use of each type of service. In general, a complete audio program (what is presented to the listener over the set of loudspeakers) may consist of a main audio service, an associated audio service which is a complete mix, or a main audio service combined with an associated audio service. The capability to simultaneously decode one main service and one associated service is required in order to form a complete audio program in certain service combinations described in this Section. This capability may not exist in some receivers.

6.2 Summary of service types

The audio service types are listed in Table 2.

Table 2 Audio Service Types

bsmod	Type of service
000 (0)	Main audio service: complete main (CM)
001 (1)	Main audio service: music and effects (ME)
010 (2)	Associated service: visually impaired (VI)
011 (3)	Associated service: hearing impaired (HI)
100 (4)	Associated service: dialogue (D)
101 (5)	Associated service: commentary (C)
110 (6)	Associated service: emergency (E)
111 (7)	Associated service: voice-over (VO)

6.3 Complete main audio service (CM)

The CM type of main audio service contains a complete audio program (complete with dialogue, music, and effects). This is the type of audio service normally provided.

The CM service may contain from 1 to 5.1 audio channels. The CM service may be further enhanced by means of the VI, HI, C, E, or VO associated services described below. Audio in multiple languages may be provided by supplying multiple CM services, each in a different language.

6.4 Main audio service, music and effects (ME)

The ME type of main audio service contains the music and effects of an audio program, but not the dialogue for the program. The ME service may contain from 1 to 5.1 audio channels. The primary program dialogue is missing and (if any exists) is supplied by simultaneously encoding a D associated service. Multiple D associated services in different languages may be associated with a single ME service.

6.5 Visually impaired (VI)

The VI associated service typically contains a narrative description of the visual program content. In this case, the VI service shall be a single audio channel. The simultaneous reproduction of both the VI associated service and the CM main audio service allows the visually impaired user to enjoy the main multi-channel audio program, as well as to follow (by ear) the on-screen activity.

The dynamic range control signal in this type of VI service is intended to be used by the audio decoder to modify the level of the main audio program. Thus the level of the main audio service will be under the control of the VI service provider, and the provider may signal the decoder (by altering the dynamic range control words embedded in the VI audio elementary stream) to reduce the level of the main audio service by up to 24 dB in order to assure that the narrative description is intelligible.

Besides providing the VI service as a single narrative channel, the VI service may be provided as a complete program mix containing music, effects, dialogue, and the narration. In this case, the service may be coded using any number of channels (up to 5.1), and the dynamic range control signal applies only to this service. The fact that the service is a complete mix shall be indicated in the AC-3 descriptor (see A/52, Annex A).

6.6 Hearing impaired (HI)

The HI associated service typically contains only dialogue which is intended to be reproduced simultaneously with the CM service. In this case, the HI service shall be a single audio channel. This dialogue may have been processed for improved intelligibility by hearing impaired listeners. Simultaneous reproduction of both the CM and HI services allows the hearing impaired listener to hear a mix of the CM and HI services in order to emphasize the dialogue while still providing some music and effects.

Besides providing the HI service as a single dialogue channel, the HI service may be provided as a complete program mix containing music, effects, and dialogue with enhanced intelligibility. In this case, the service may be coded using any number of channels (up to 5.1). The fact that the service is a complete mix shall be indicated in the AC-3 descriptor (see A/52, Annex A).

6.7 Dialogue (D)

The D associated service contains program dialogue intended for use with an ME main audio service. The language of the D service is indicated in the AC-3 bit stream, and in the audio descriptor. A complete audio program is formed by simultaneously decoding the D service and the ME service and mixing the D service into the center channel of the ME main service (with which it is associated).

If the ME main audio service contains more than two audio channels, the D service shall be monophonic (1/0 mode). If the main audio service contains two channels, the D service may also contain two channels (2/0 mode). In this case, a complete audio program is formed by simultaneously decoding the D service and the ME service, mixing the left channel of the ME service with the left channel of the D service, and mixing the right channel of the ME service with the right channel of the D service. The result will be a two channel stereo signal containing music, effects, and dialogue.

Audio in multiple languages may be provided by supplying multiple D services (each in a different language) along with a single ME service. This is more efficient than providing multiple CM services, but, in the case of more than two audio channels in the ME service, requires that dialogue be restricted to the center channel.

Some receivers may not have the capability to simultaneously decode an ME and a D service.

6.8 Commentary (C)

The commentary associated service is similar to the D service, except that instead of conveying essential program dialogue, the C service conveys optional program commentary. The C service may be a single audio channel containing only the commentary content. In this case, simultaneous reproduction of a C service and a CM service will allow the listener to hear the added program commentary.

The dynamic range control signal in the single channel C service is intended to be used by the audio decoder to modify the level of the main audio program. Thus the level of the main audio service will be under the control of the C service provider, and the provider may signal the decoder (by altering the dynamic range control words embedded in the C audio elementary stream) to reduce the level of the main audio service by up to 24 dB in order to assure that the commentary is intelligible.

Besides providing the C service as a single commentary channel, the C service may be provided as a complete program mix containing music, effects, dialogue, and the commentary. In this case the service may be provided using any number of channels (up to 5.1). The fact that the service is a complete mix shall be indicated in the AC-3 descriptor (see A/52, Annex A).

6.9 Emergency (E)

The E associated service is intended to allow the insertion of emergency or high priority announcements. The E service is always a single audio channel. An E service is given priority in transport and in audio decoding. Whenever the E service is present, it will be delivered to the audio decoder. Whenever the audio decoder receives an E type associated service, it will stop reproducing any main service being received and only reproduce the E service out of the center channel (or left and right channels if a center loudspeaker does not exits). The E service may also be used for non-emergency applications. It may be used whenever the broadcaster wishes to force all decoders to quit reproducing the main audio program and reproduce a higher priority single audio channel.

6.10 Voice-over (VO)

The VO associated service is a single channel service intended to be reproduced along with the main audio service in the receiver. It allows typical voice-overs to be added to an already encoded audio elementary stream without requiring the audio to be decoded back to baseband and then re-encoded. It is always a single audio channel. It has second priority (only the E service has higher priority). It is intended to be simultaneously decoded and mixed into the center channel of the main audio service. The dynamic range control signal in the VO service is intended to be used by the audio decoder to modify the level of the main audio program. Thus the level of the main audio service may be controlled by the broadcaster, and the broadcaster may signal the decoder (by altering the dynamic range control words embedded in the VO audio elementary stream) to reduce the level of the main audio service by up to 24 dB during the voice-over.

Some receivers may not have the capability to simultaneously decode and reproduce a voice-over service along with a program audio service.

7. AUDIO ENCODER INTERFACES

7.1 Audio encoder input characteristics

Audio signals which are input to the digital television system may be in analog or digital form. Audio signals should have any DC offset removed before being encoded. If the audio encoder does not include a DC blocking high pass filter, the audio signals should be high pass filtered before being applied to the encoder. In general, input signals should be quantized to at least 16-bit resolution. The audio compression system can convey audio signals with up to 24-bit resolution. Physical interfaces for the audio inputs to the encoder may be defined as voluntary industry standards by the AES, SMPTE, or other standards organizations.

7.2 Audio encoder output characteristics

Conceptually, the output of the audio encoder is an elementary stream which is formed into PES packets within the transport subsystem. It is possible that systems will be implemented wherein the formation of audio PES packets takes place within the audio encoder. In this case, the output(s) of the audio encoder(s) would be PES packets. Physical interfaces for these outputs (elementary streams and/or PES packets) may be defined as voluntary industry standards by SMPTE or other standards organizations.

ANNEX C

(Normative)

SERVICE MULTIPLEX AND TRANSPORT SYSTEMS CHARACTERISTICS

1. SCOPE

This Annex describes the transport layer characteristics and normative specifications of the Digital Television Standard.

2. NORMATIVE REFERENCES

The following documents contain provisions which in whole or in part, through reference in this text, constitute provisions of this Standard. At the time of publication, the editions indicated were valid. All standards are subject to revision and amendment, and parties to agreements based on this Standard are encouraged to investigate the possibility of applying the most recent editions of the documents listed below.

ATSC Standard A/52 (1995), *Digital Audio Compression (AC-3)*.

ISO/IEC IS 13818-1, International Standard (1994), *MPEG-2 Systems*.

ISO/IEC IS 13818-2, International Standard (1994), *MPEG-2 Video*.

ISO/IEC CD 13818-4, MPEG Committee Draft (1994), *MPEG-2 Compliance*.

The normative reference for the Program Guide will be the standard developed from ATSC document T3/S8-050, "Program Guide for Digital Television".

The normative reference for System Information will be the standard developed from ATSC document T3/S8-079, "System Information for Digital Television".

3. COMPLIANCE NOTATION

As used in this document, "*shall*" or "*will*" denotes a mandatory provision of the standard. "*Should*" denotes a provision that is recommended but not mandatory. "*May*" denotes a feature whose presence does not preclude compliance, that may or may not be present at the option of the implementor.

4. SYSTEM OVERVIEW

The transport format and protocol for the Digital Television Standard is a compatible subset of the MPEG-2 Systems specification defined in ISO/IEC 13818-1. It is based on a fixed-length packet transport stream approach which has been defined and optimized for digital television delivery applications.

As illustrated in Figure 1, the transport function resides between the application (e.g., audio or video) encoding and decoding functions and the transmission subsystem. The encoder's transport subsystem is responsible for formatting the coded elementary streams and multiplexing the different components of the program for transmission. At the receiver, it is responsible for recovering the elementary streams for the individual application decoders and for the corresponding error signaling. The transport subsystem also incorporates other higher protocol layer functionality related to synchronization of the receiver.

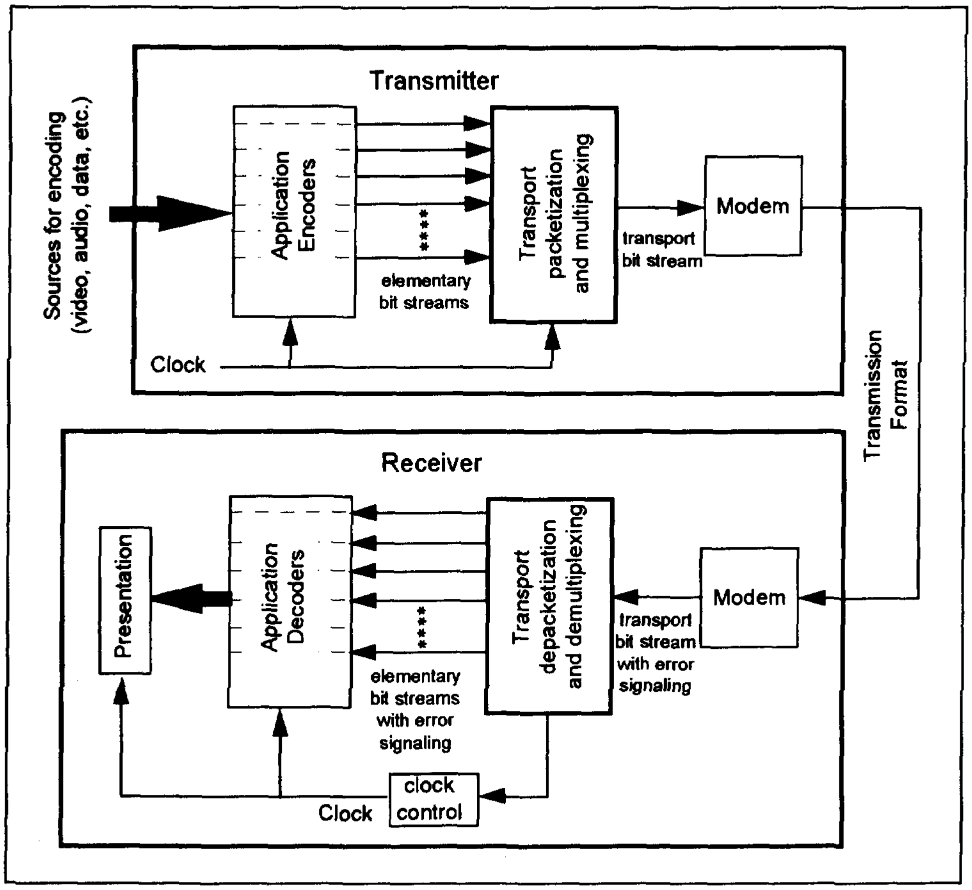

Figure 1. Sample organization of functionality in a transmitter-receiver pair for a single program.

The overall system multiplexing approach can be thought of as a combination of multiplexing at two different layers. In the first layer, single program transport bit streams are formed by multiplexing transport packets from one or more Packetized Elementary Stream (PES) sources. In the second layer, many single program transport bit streams are combined to form a system of programs. The Program Specific Information (PSI) streams contain the information relating to the identification of programs and the components of each program.

Not shown explicitly in Figure 1, but essential to the practical implementation of this Standard, is a control system that manages the transfer and processing of the elementary streams from the application encoders. The rules followed by this control system are not a part of this Standard but must be established as recommended practices by the users of the Standard. The control system implementation shall adhere to the requirements of the MPEG-2 transport system as specified in ISO/IEC 13818-1 with the additional constraints specified in this Standard. These constraints may go beyond the constraints imposed by the application encoders.

5. SPECIFICATION

This Section constitutes the normative specification for the transport system of the Digital Television Standard. The syntax and semantics of the specification conform to ISO/IEC 13818-1 subject to the constraints and conditions specified in this Standard. This Section of the Standard describes the coding constraints that apply to the use of the MPEG-2 systems specification in the digital television system.

5.1 MPEG-2 Systems standard

The transport system is based on the transport stream definition of the MPEG-2 Systems standard as specified in ISO/IEC 13818-1.

5.1.1 Video T-STD

The video T-STD is specified in Section 2.4.2.3 of ISO/IEC 13818-1 and follows the constraints for the level encoded in the video elementary stream.

5.1.2 Audio T-STD

The audio T-STD is specified in Section 3.6 of Annex A of ATSC Standard A/52.

5.2 Registration descriptor

This Standard uses the registration descriptor described in Section 2.6.8 of ISO/IEC 13818-1 to identify the contents of programs and elementary streams to decoding equipment.

5.2.1 Program identifier

Programs which conform to this specification will be identified by the 32-bit identifier in the section of the Program Map Table (PMT) detailed in Section 2.4.4.8 of ISO/IEC 13818-1. The identifier will be coded according to Section 2.6.8, and shall have a value of 0x4741 3934.

5.2.2 Audio elementary stream identifier

Audio elementary streams which conform to this specification will be identified by the 32-bit identifier in the section of the Program Map Table (PMT) detailed in Section 2.4.4.8 of ISO/IEC 13818-1. The identifier will be coded according to Section 2.6.8, and shall have a value of 0x4143 2D33.

5.3 The program paradigm

The program paradigm specifies the method that shall be used for allocating the values of the Packet Identifier (PID) field of the transport packet header in a systematic manner. Within one transport multiplex, television programs that follow the program paradigm are assigned a program number ranging from 1 to 255. The binary value of the program number is used to form b_{11} through b_4 of the PID. Programs adhering to the paradigm shall have b_{12} equal to '0'. Programs not adhering to the paradigm shall have b_{12} equal to '1'.

We further define:

- base_PID = program number << 4

where program number refers to each program within one transport multiplex and corresponds to the 16-bit program_number identified in PAT and PMT.

The b_0 through b_3 of the PID are assigned according to Table 1.

The paradigm to identify the transport bit streams containing certain elements of the program is defined in Table 1.

Table 1 PID Assignment for the Constituent Elementary Streams of a Program

Name	PID Definition	Description
PMT_PID	base_PID+0x0000	PID for the bit stream containing the program_map_table for the program.
Video_PID	base_PID+0x0001	PID for the bit stream containing the video for the program.
PCR_PID	base_PID+0x0001	Implies the video bit stream also carries the PCR values for the program
Audio_PID	base_PID+0x0004	PID for the bit stream containing the primary audio for the program. The primary audio shall be a complete main audio service (CM) as defined by ATSC Standard A/52 and shall contain the complete primary audio of the program including all required voice-overs and emergency messages.
Data_PID	base_PID+0x000A	PID for the bit stream containing the data for the program.

The program_map_table must be decoded to obtain the PIDs for services not defined by the paradigm but included within the program (such as a second data channel). According to the program paradigm, every 16th PID is a PMT_PID and may be assigned to a program. If a PMT_PID is assigned to a program by the program paradigm, the next 15 PIDs after that PMT_PID are reserved for elements of that program and shall not be otherwise assigned.

5.4 Constraints on PSI

The program constituents for all programs, including television programs that follow the program paradigm and other programs or services that do not follow the program paradigm, are described in the PSI. There are the following constraints on the PSI information:

- Only one program is described in a PSI transport bit stream corresponding to a particular PMT_PID value. A transport bit stream containing a program_map_table shall not be used to transmit any other kind of PSI table (identified by a different table_id).
- The maximum spacing between occurrences of a program_map_table containing television program information shall be 400 ms.
- The program numbers are associated with the corresponding PMT_PIDs in the PID0 Program Association Table. The maximum spacing between occurrences of section 0 of the program_association_table is 100 ms.
- The video elementary stream section shall contain the Data stream alignment descriptor described in Section 2.6.10 of ISO/IEC 13818-1. The alignment_type field shown in Table 2-47 of ISO/IEC 13818-1 shall be 0x02.
- Adaptation headers shall not occur in transport packets of the PMT_PID for purposes other than for signaling with the discontinuity_indictor that the version_number (Section 2.4.4.5 of ISO/IEC 13818-1) may be discontinuous.
- Adaptation headers shall not occur in transport packets of the PAT_PID for purposes other than for signaling with the discontinuity_indicator that the version_number (Section 2.4.4.5 of ISO/IEC 13818-1) may be discontinuous.

5.5 PES constraints

Packetized Elementary Stream syntax and semantics shall be used to encapsulate the audio and video elementary stream information. The Packetized Elementary Stream syntax is used to convey the Presentation Time-Stamp (PTS) and Decoding Time-Stamp (DTS) information required for decoding audio and video information with synchronism. This Section describes the coding constraints for this system layer.

Within the PES packet header, the following restrictions apply:

- PES_scrambling_control shall be coded as '00'.
- ESCR_flag shall be coded as '0'.
- ES_rate_flag shall be coded as '0'.
- PES_CRC_flag shall be coded as '0'.

Within the PES packet extension, the following restrictions apply.

- PES_private_data_flag shall be coded as '0'.
- pack_header_field_flag shall be coded as '0'.
- program_packet_sequence_counter_flag shall be coded as '0'.
- P-STD_buffer_flag shall be coded as '0'.